# BEBR

FACULTY WORKING PAPER NO. 1402



## An Analysis of Nonlinearities, Heteroskedasticity, and Functional Form in the Market Model

*Bill McDonald*
*Cheng-Few Lee*



College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

# BEBR

FACULTY WORKING PAPER NO. 1402

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

October 1987

Bill McDonald
University of Notre Dame

Cheng-Few Lee, Professor
Department of Finance

An Analysis of Nonlinearities, Heteroskedasticity, and Functional Form in the Market Model

## Abstract

Using a generalized specification of the single-index market model, this study examines the sources of statistical anomalies previously found in estimating the market model. Two generalized models are developed for juxtaposition with the traditional linear specification. The most general model is a Box-Cox model with different $\lambda$'s and heteroskedastic errors. The empirical results indicate that previous findings of significant "nonlinearities" are primarily attributable to nonnormalities and unequal variance.


KEY WORDS: Box-Cox model, Nonnormalities, Asset pricing

# 1. INTRODUCTION

The return generating process serving as the foundation of most empirical tests of asset pricing is the single-index market model, expressed for a given security in ex-post terms as:

$$R_t = \alpha + \beta R_{mt} + \varepsilon_t \quad , \tag{1}$$

where:

$R_t$ = the security's return in time t minus the return on a T-bill in the same period,

$R_{mt}$ = the value-weighted market return in time t minus the return on a T-bill in the same period,

$\alpha, \beta$ = parameters to be estimated, and

$\varepsilon_t \sim N(0, \sigma^2)$ .

Although empirical tests of the capital asset pricing model (CAPM) are subject to Roll's (1977) qualifications, the market model remains an important specification of the returns generating process, whether used explicitly in CAPM tests or simply as an adjustment to examine returns around an economic event.

Using flexible functional form techniques, a number of studies (e.g., Lee 1976a, Lee 1976b, and McDonald 1983) document the presence of significant departures from a linear model with normally distributed errors. In some cases, the authors go on to label their findings as evidence of significant "nonlinearities" in the returns model. Although this conclusion is statistically valid, the sources of the nonlinearities and their implications have not been completely analyzed.

Nonlinearities reported in other work can be attributed to three possible sources: 1) nonnormalities in the residuals, 2) a heteroskedastic error variance, or 3) a nonlinear (nonadditive) relationship between the variables. Given the significance of the linearity assumption in asset pricing theories, it is important to determine the actual source of departures from the basic market model by partitioning these effects.

This study uses a series of specifications including generalized functional forms and a heteroskedastic error model to address distributional and linearity issues that have surfaced in previous research. Three different models are estimated for each of 1,164 securities using monthly security returns. The first model is the traditional linear market model. The second is a generalized Box-Cox model with different transformation parameters for the dependent and independent variables. The third extends the generalized Box-Cox model to allow for conditional heteroskedasticity. The residuals from the linear model exhibit statistically significant skewness and excess kurtosis, suggesting a possible source of significant transformations. In the Box-Cox models, the transformation parameter for the independent variable is usually not statistically significant. The transformation for the dependent variable is usually significant, and it is shown that its significance appears to be attributable to the skewness of residuals in the linear model. Also, there is evidence of conditional heteroskedasticity, which appears to be the cause of excess kurtosis of the residuals from the linear model.

The econometric models to be tested are derived sequentially in Section II. References to previous studies and theoretical considerations will be incorporated in the model development. The third section details the estimation procedure and data base. Section IV of the paper presents the empirical results. The final section presents a compendium of the pertinent findings.

## II. MODEL DEVELOPMENT

In this section the market model of equation (1) is extended first to incorporate a generalized functional form, and then a combined heteroskedastic/functional form specification.

The linear functional form of the market model in equation (1) can be generalized by applying the transformation of Box and Cox (1964), as:

$$R_t'^{(\lambda 1)} = \alpha + \beta R_{mt}'^{(\lambda 2)} + \varepsilon_t \tag{2.1}$$

where:

$$X^{(\lambda)} = \begin{cases} (X^{\lambda}-1) / \lambda & \lambda \neq 0 \\ \ln X & \lambda = 0 , \end{cases} \tag{2.2}$$

and returns are expressed in wealth relative form (e.g., R'=1+R). Applying the transformation in this manner, the model reduces to a linear form when $\lambda_1=\lambda_2=1$ and a logarithmic form when $\lambda_1=\lambda_2=0$. The parameters of the model are estimated using maximum likelihood techniques with the concentrated log-likelihood function given by:

$$\ln L = -T \ln \sigma + -1/(2\sigma^2) \ \Sigma \, \varepsilon^2_t + J \ , \tag{2.3}$$

where J is the Jacobian of the inverse transformation of $R'_t{}^{(\lambda_1)}$ to $R'_t$ given by:

$$(\lambda_1-1) \ \Sigma \ln R'_t \ . \tag{2.4}$$

It is assumed in equation (2.2) that for some values of $\lambda_1$ and $\lambda_2$ the transformed

observations will be normally distributed with constant variance, and conditional expectation that is linear in the transformed independent variable. Thus, the estimated transformation parameters will, to the extent possible, correct for nonlinearities, unequal variance, and nonnormalities. Each of these characteristics has a crucial role in the theoretical and econometric application of the market model. Unfortunately, in application of the GFF method, these aberrations are corrected simultaneously and the method provides no specific interpretation of the estimated transformation. Although no exact method has been proposed to isolate the various effects of the transformations, the method developed in this study attempts to disaggregate the effects using a series of tests.

Asymptotic standard errors for the parameter estimates are given by the Cramer-Rao lower bound. Both Hinkley and Runger (1984) and Spitzer (1984) show that the estimated standard errors in a Box-Cox model are sensitive to the scale of the variables; however, this issue is not a problem in this case since the variables are universally expressed as returns.

Although previous studies have applied the GFF methodology, none have attempted to isolate the econometric effects so that the estimated transformation can be explicitly interpreted. The majority of previous studies applying the transformation have also restricted the GFF model by assuming a single transformation across all variables (i.e., $\lambda_1=\lambda_2$). The more general specification, where the security returns and market returns have distinct transformation parameters, has a number of important implications. As suggested by Box and Tidwell (1962) and Zarembka (1974), the transformations on dependent versus independent variables tend to focus on different issues. The transformation of an independent variable is concentrating on the linearity issue, since the independent

variables are conditioning the value of the dependent variable. The transformation of the random dependent variable focuses on the error term assumptions of normality and constant variance. Thus separate transformations should allow the various factors effecting the estimated $\lambda$'s to be partitioned to determine if the significant estimates observed in previous studies are attributable to nonlinearities, heteroskedasticity, or nonnormalities. For the single transformation case (i.e., $\lambda_1=\lambda_2$), Zarembka (1974) delineates the effects of nonnormalities and heteroskedasticity on the estimate of $\lambda$. The presence of nonnormalities is addressed in this study by monitoring distributional measures of the error terms in the empirical models.

Substantial evidence of heteroskedasticity in the market model is provided by a number of studies (e.g., Bey and Pinches 1980, Giaccotto and Ali 1982, or McDonald and Morris 1983). From the results of previous studies, the finding of heteroskedasticity appears to depend upon the time period studied. For a time interval overlapping the one used in this study, Bey and Pinches find significant heteroskedasticity in approximately 45 percent of the securities sampled. Using a basic heteroskedastic specification, Lahiri and Egy (1981) test a model that provides for the joint estimation of functional form under conditions of unequal variance. Their results confirm the bias in estimating $\lambda$ in the presence of heteroskedasticity and emphasize the importance of simultaneous testing of functional form and heteroskedasticity. A reformulation similar to theirs will be applied to the GFF model of equation (2.1).

Consistent with previous studies, the variance is assumed to be a function of $R_{mt}$ and is specified as

$$\sigma^2_t = \delta_0 + \delta_1 R_{mt} + \delta_2 R^2_{mt} \tag{2.5}$$

This specification is general enough to include a random coefficients process when $\delta_1$ is equal to zero. Given the variance structure of equation (2.5), the likelihood function for the heteroskedastic specification is:

$$\ln L = -\Sigma \ln \sigma_t - (1/2)\, \Sigma\, (\varepsilon_t / \sigma^2_t) + (\lambda_1 - 1)\, \Sigma \ln R'_t \,. \tag{2.6}$$

## III. ESTIMATION AND DATA

The parameters for the likelihood function of equation (2.3) and its extension to the heteroskedastic specification of equation (2.6) are estimated using full-information maximum likelihood techniques. Specifically, the likelihood functions are maximized using the Davidon-Fletcher-Powell algorithm contained in the numerical optimization software "GQOPT," documented extensively in Goldfeld and Quandt (1976). The function provides for an error return in the optimization for cases where the likelihood function is undefined (in this case, where the variance is negative). In estimating the likelihood functions, the optimizations were generally well-behaved and converged quite rapidly.

Data for the study were taken from the Center for Research in Security Prices' monthly returns tape for the period of January 1973 to December 1979. This particular period was selected to parallel the results of McDonald (1983) and because it is a period that has been shown to exhibit relatively high levels of

heteroskedasticity (see Bey and Pinches 1980). A total of 1,164 securities had complete information for the 84 month time interval. The market return was measured using the value-weighted index reported on the CRSP tapes.

## IV. RESULTS

### 4.1 Comparison of the Model Specifications

The importance of extending the linear model to the more general specifications is examined using the likelihood ratio test. Under general conditions $-2(\ln L_1 - \ln L_2)$ is distributed as a $\chi^2_{(r)}$, where $L_1$ is the likelihood of the constrained model, $L_2$ is the unconstrained likelihood value, and r is the number of parameters for which $L_1$ specifies given values. The percentage of securities where one form provides a significant improvement over a corresponding restricted form is shown in Table 1. (All statistical tests are at the $\alpha$=0.05 level unless otherwise stated.)

Insert Table 1 about here

In comparison to the linear market model, the GFF model is superior in more than 46 percent of the cases. Whether this result is attributable to corrections for nonnormality, heteroskedasticity, or true nonlinearities is not determinable at this point. The HET-GFF specification provides improvement beyond the GFF model in approximately 47 percent of the cases. Thus the more general forms merit additional attention given the observed frequency with which they contribute significant information.

## 4.2 Heteroskedasticity and the Market Model

The purpose of including heteroskedasticity in the model specifications is to determine the impact of unequal variances on the distribution of the market model residuals and the effect of unequal variance on estimates of the transformations. Evidence of significant heteroskedasticity in comparing the HET-GFF and GFF models (from Table 1) is consistent with the findings of Bey and Pinches (1980), suggesting that previous findings of heteroskedasticity are not merely a result of functional form misspecification. Significance tests on the estimates of $\delta_1$ and $\delta_2$ provide some indication of the characteristics underlying the heteroskedasticity. For the securities where $\delta_1$ or $\delta_2$ was significant, $\delta_2$ was singly significant in 53 percent of the cases, $\delta_1$ was singly significant in 33 percent of the cases, with both coefficients significant in the remaining 14 percent of the cases. This finding indicates that the variance expressed as a function of $R^2{}_{mt}$ is a predominant form, as found by Giaccotto and Ali (1982); however, other forms cannot be completely excluded.

## 4.3 Nonnormalities and the Market Model Specification

To determine the effects of the various market model specifications on the distributional properties of the underlying process, a measure of skewness and kurtosis was calculated from the market model residuals of each case. The average value for these statistics for the 1,164 cases is reported in Table 2.

---

Insert Table 2 about here

---

Significance tests of the distributional measures based on aggregates of individual tests are not appropriate in this case because of the cross-correlation in the security residuals. A crude correction for the significance tests of excess skewness can be made by using an analysis of variance approach to estimate:

$$z^3_{it} = \alpha_{1i} + \alpha_{2m} + \eta_{im} , \tag{4.3.1}$$

where $z_{it}$ is the standardized residual, $\alpha_{1i}$ represents the individual mean of the cubed residual for the ith security, $\alpha_{2m}$ represents the month effect which captures cross-security correlation, and $\eta_{im}$ is an uncorrelated error. The null hypothesis is that $\alpha_{1i}=0$ and $\alpha_{2m}=0$. Similarly, a test for excess kurtosis can be estimated from:

$$z^4_{it} = 3 + \theta_{1i} + \theta_{2m} + v_{im} . \tag{4.3.2}$$

The joint F-tests for the coefficients in equation 4.3.1 and 4.3.2 appear in Table 2 and document the presence of significant skewness and kurtosis in the residuals of the LIN model. The descriptive statistics and significance tests confirm the impact of the transformations in adjusting for nonnormalities. Since the effects of the transformation are confounded in the GFF specification, the results of the more general HET-GFF model are of primary interest.

For the HET-GFF specification where the effects of the functional form

transformations should be concentrated on correcting for nonlinearities and nonnormalities, the presence of skewness has been virtually eliminated. There is also a substantial reduction in kurtosis attributable to the heteroskedastic specification. These results indicate that the thick tails and central peakedness of the distribution of returns can be attributed to unequal variances and is not necessarily the result of a nonnormal stochastic process. These results are consistent with the findings of Rosenberg and Marathe (1979) and provide limited support for the subordinated normal hypothesis of return distributions (see Clark 1973).

## 4.4 Isolating the Effects of the Transformation Parameters

The series of models tested in this study allows the effects of the functional form parameters to be partitioned to some degree. The mean value for the estimates of $\lambda_1$ and $\lambda_2$ across all securities, along with the percent of cases where the transformations were statistically different from one are presented in Table 3. From the distribution of the estimated $\lambda$'s, there does not seem to be any apparent tendency for the transformation parameters to center around one particular value. Although, as previously noted, the heteroskedastic specification has a substantial effect on the distribution of the residuals, it does not appear to have a substantive effect on the estimated transformations.

Insert Table 3 about here

The effects of unequal variance on the estimated transformations has been

avoided by extending the model to allow for heteroskedasticity. The effects of nonnormalities and nonlinearities, however, are still confounded. Unfortunately there is no exact means of disaggregating these effects. A series of tests can be applied to provide evidence of the causes of significant transformations.

First note, as suggested by Box and Tidwell (1962), that the transformation of the independent variable concentrates on additivity of effect. The relatively small proportion of significant $\lambda_2$'s reported in Table 3 provides evidence supporting the linearity hypothesis and is consistent with the findings of McDonald (1983).

Second, from Table 2 the average skewness drops from .525 to -.001 in the GFF model. Given that the significance $\lambda_1$ is affected by both nonlinearities and nonnormalities, the skewness results would suggest that much of the significance is attributable to corrections for nonnormalities. To test this proposition all securities with significant skewness in the market model residuals are eliminated from the sample. In this case, the occurrence of significant $\lambda_1$'s is no more than expected by chance.

In summary, the accumulation of evidence suggests that the Box-Cox transformation reduces skewness in the dependent variable. In the absence of skewness, the transformation parameter for the dependent variable is no longer significant. After adjusting for heteroskedasticity and nonnormalities, the transformation parameters are generally not significant. The "nonlinearities" identified in previous studies using the GFF methodology (e.g., Lee 1976a), can be attributed to distributional aberrations and not the the functional relationship between the market model variables.

# V. CONCLUSIONS

Using a sample of 1,164 securities, this study examines the theoretical and statistical implications of certain econometric phenomena occurring in the estimation of the market model. The empirical results have significant implications for the asset pricing theories underlying the single-factor market model and their common basis of linearity. In summary, the empirical results from the sample tested in this study suggest that:

1. The parameters of the generalized market models are statistically significant in a substantial number of cases. Specifically, for more than half the securities tested the parameters $\lambda_1$ and $\lambda_2$ of the generalized Box-Cox model were jointly statistically significant as were the parameters of the model for heteroskedasticity.
2. The residuals from estimates of the linear market model exhibit significant nonnormalities. Correcting for unequal variances reduces the level of kurtosis substantially.
3. The statistical significance of the transformation parameter for the dependent variable is primarily attributable to skewness in the market model residuals. Transformation of the dependent variable eliminates the skewness in the market model residuals.

**Acknowledgments:** Data for the study were made available through the Jesse H. Jones Research Data Base at the University of Notre Dame. Financial support was provided in part by the Herrick Foundation Fund. The manuscipt has benefitted from the substantive comments of an Associate Editor and two anonymous reviewers

## REFERENCES

Bey, R.P. and Pinches, G.E. (1980), "Additional Evidence of Heteroskedasticity in the Market Model," Journal of Financial and Quantitative Analysis, 15, 299-322.

Box, G.E.P. and Cox, D.R. (1964), "An Analysis of Transformations," Journal of the Royal Statistical Society, 26B, 211-243.

Box, G.E.P. and Tidwell, P.W. (1962), "Transformation of the Independent Variables," Technometrics, 4, 531-550.

Clark, P.K., (1973), "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," Econometrica, 41, 135-159.

Giaccotto, C. and Ali, M. (1982) "Optimum Distribution-Free Tests and Further Evidence of Heteroscedasticity in the Market Model," *Journal of Finance*, 37, 1247-1258.

Goldfeld, S.M. and Quandt, R.E. (1976), Nonlinear Methods in Econometrics, Amsterdam, North-Holland Publishing Company.

Hinkley, D.V., and Runger, G. (1984), "The Analysis of Transformed Data," Journal of the American Statistical Association, 79, 302-312.

Lahiri, K. and Egy, D. (1981), "Joint Estimation and Testing for Functional Form and Heteroskedasticity," Journal of Econometrics, 15, 299-307.

Lee, C.F. (1976a), "Functional Form and the Dividend Effect in the Electric Utility Industry," Journal of Finance, 31, 1481-1486.

________ (1976b), "Investment Horizon and the Functional Form of the Capital Asset Pricing Model," Review of Economics and Statistics, 58, 356-363.

McDonald, B. (1983), "Functional Forms and the Capital Asset Pricing Model," The Journal of Financial and Quantitative Analysis, 18, 319-329.

McDonald, B. and Morris, M. (1983), "The Existence of Heteroscedasticity and Its Effects on Estimates of the Market Model Parameters," The Journal of Financial Research, 6, 115-126.

Roll, R. (1977), "A Critique of the Asset Pricing Theory's Tests; Part 1: On Past and Potential Testability of the Theory," Journal of Financial Economics, 4, 129-175.

Rosenberg, B. and Marathe, V. (1979), "Tests of Capital Asset Pricing Hypotheses," in H. Levy, ed., Research in Finance, Connecticut, JAI Press, 115-223.

Spitzer, J.J. (1984), "Variance Estimates in Models with the Box-Cox Transformation: Implications for Estimation and Hypothesis Testing," The Review of Economics and Statistics, 66, 645-652.

Zarembka, P. (1974), "Transformation of Variables in Econometrics," in P. Zarembka, ed., Frontiers in Econometrics, New York, Academic Press, 81-104.

**Table 1**

Specification Comparisons:
Percentage of Securities with Significant
Difference between the General and Restricted
Market Model Forms

| Generalized Form | Restricted Counterpart LIN | GFF |
|---|---|---|
| GFF | 46.8% | ----- |
| HET-GFF | 67.8 | 47.7% |

NOTE: Tests are performed for each of the 1,164 securities using the likelihood ratio test with $\alpha$=0.05.

**Table 2**

Skewness, Kurtosis, and F-tests of Significance for the Three Empirical Models[a]

| Model | Skewness: Mean (Std.Dev.) | Skewness: F-test[b] $\alpha_{1i}=\alpha_{2m}=0$ (p-level) | Kurtosis: Mean (Std.Dev.) | Kurtosis: F-test[c] $\theta_{1i}=\theta_{2m}=0$ (p-level) |
|---|---|---|---|---|
| LIN | 0.525 (0.639) | 1.13 (0.002) | 2.109 (3.035) | 1.27 (0.001) |
| GFF | -0.001 (0.275) | 0.42 (0.999) | 0.808 (1.167) | 3.99 (0.001) |
| HET-GFF | -0.006 (0.145) | 0.34 (0.999) | 0.417 (0.861) | 0.78 (0.999) |

[a]The skewness and kurtosis measures are estimated from the residuals of each specification for each security. The mean and standard deviation across the 1,164 securities for each statistic is reported in the table. Calculations of skewness and kurtosis are based on Fisher's k-statistics.

[b]Joint F-test for excess skewness with 1,174 degrees of freedom in the numerator and 96,601 degrees of freedom in the denominator. See equation (4.3.1).

[c]Joint F-test for excess kurtosis with 1,174 degrees of freedom in the numerator and 96,601 degrees of freedom in the denominator. See equation (4.3.2).

**Table 3**

Estimated Transformation Parameters

| Model | Mean (Std. Dev.) | $H_0$: $\lambda_1=1$ Percent Significant | Mean (Std. Dev.) | $H_0$: $\lambda_2=1$ Percent Significant |
|---|---|---|---|---|
| GFF | -0.451 (1.186) | 48.9% | -1.201 (5.177) | 12.5% |
| HET-GFF | -0.462 (1.042) | 44.4 | -1.042 (5.557) | 19.8 |